UNEMPLOYMENT ISSUES IN BARAOLT REGION

Assist. Ph.D Kulcsár Erika
University "Babeş-Bolyai"
Faculty of Economics and Business Administrations
Cluj Napoca, Romania
Economist Sütő Gábor

Abstract: Many methods of multivariate analysis are based on metric variables. In the case of questionnaire-based surveys qualitative data usually prevail, measured nominally or ordinal. This paper contains the discriminant analysis from data recorded from a sample of 100 respondents, with regard to: whether they have a job or not, number of family members and the importance of having a job according to the specialization graduated.

JEL classification: C32, M31

Key words: variable, ANOVA, level of significance, the discriminant function, Wilks λ

Introduction

The discriminant analysis based on primary data obtained from carrying a quantitative marketing research in Baraolt region. The basic feature of these methods consists of the existence of one or more dependent variables and several independent variables.

In the case of the discriminant analysis the following categories of variable occur:

- a nominal dependent variable measured with two categories (NO, YES)
- 2 independent variables: one representing a proportional scale and an interval scale.

Through the discriminant analysis we can determine what distinguishes persons who have a job from those who do not have a job according to the number of family members and the importance attributed to accepting a job according to the specializations graduated.

Processing of the primary date – Discriminant analysis

The sample was composed of 100 persons more precisely the questionnaires were completed by 50 people without a job and by 50 people who work. As shown in the table no.1 those who responded to the variables that make up the starting point of the discriminant analysis are 82 respondents, 39 unemployed and 43 employed.

Although in our opinion it would have been more logical that the average number of unemployed family members to be smaller we got just the opposite and the number of family members for the working respondents is 2.83 (3 persons) while with the unemployed respondents it is 4.38 (4 persons). For the unemployed persons it is more important to find a job according to the specialty graduated than for those who

have a jobs. Although we expected that the values obtained to show the opposite results.

The statistical significance was tested using the variance (ANOVA)

Table no. 1 Averages of independent variables in the groups

Group Statistics

Do you have a job?		Mean	Std. Deviation	Valid N (listwise) Unweighted	Valid N (listwise) Weighted
No	How many members does your family have?	4.38	1.016	39	39.000
	According to the gratuated specialization	3.54	1.570	39	39.000
Yes	How many members does your family have?	2.86	.990	43	43.000
	According to the gratuated specialization	2.37	1.648	43	43.000
Total	How many members does your family have?	3.59	1.257	82	82.000
	According to the gratuated specialization	2.93	1.705	82	82.000

Table no. 2 ANOVA tests results for equal averages

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
How many members does your family have?	.629	47.276	1	80	.000
According to the gratuated specialization	.882	10.714	1	80	.002

F values in the table no.2 have higher values, being higher than the values obtained from the table of the Fisher's distribution law according to these degrees of freedom. This is highlighted by the minimum significance levels that can accept alternative hypothesis corresponding to the two variables, which are smaller than 0.05.

Therefore, both variables have a significant discriminating power. "Number of family members" has a greater discriminating power as the coefficient λ (Wilks' Lambda) is lower and the value F is greater than for the variable "According to specialization graduated".

In the table no.3 the correlation coefficient between independent variables is calculated. Analyzing the obtained values we can see that the conditions that the independent variables are not to be collinear is respected.

Since the dependent variable has only two categories a single discriminant function is obtained. The table below (no.4) reproduced the value of the discriminant function ("Eigenvalue"), representing the sum of squares between groups and sum of squares within the groups at the level of the discriminant function, this value it is 0.872.

Table no. 3 Correlation between independent variables

Pooled Within-Groups Matrices

		How many members does your family have?	According to the gratuated specialization
Correlation	How many members does your family have?	1.000	-.200
	According to the gratuated specialization	-.200	1.000

Table no. 4 The value of the discriminant function

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.872[a]	100.0	100.0	.683

a. First 1 canonical discriminant functions were used in the analysis.

The testing of the significance of the discriminant function power is made based on Wilks λ coefficient, which must have a value as close to zero. [25]

Table no. 5 Testing the discriminant function significance.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.534	49.546	2	.000

As outlined in the table no. 5, the value $\chi^2_{calc} = 49.546 > \chi^2_{0.05;2} = 5.99$ so the alternative hypothesis is accepted according to which the function obtained has a significant discriminatory power. This is also observed in the level of minimum significance value that can accept the alternative hypothesis: 0.005<0.05. Since the discriminant function is statistically significant, it can be built on the coefficients corresponding to the independent variables and on the constant term in the table no.6.

Table no. 6 Discriminant function coefficients

Canonical Discriminant Function Coefficients

	Function
	1
How many members does your family have?	.937
According to the gratuated specialization	.360
(Constant)	-4.411

Unstandardized coefficients

[25] Cristinel C. - Marketing information system. Marketing data analysis and processing. Applications in SPSS, Publisher Infomarket, Braşov, 2006, p. 243

The discriminant function is:

$y = -4.411+0.937 \cdot x_1 + 0.360 \cdot x_2$

where:

y = discriminant function („Function 1")

x_1 - independent variable 1 („Number of family members")

x_2 - independent variable 2 („According to the graduated specialization")

The coefficients can be converted to standardized values, corresponding to the values of the normal distribution law.

Table no. 7 Standardized coefficients of discriminant function

tandardized Canonical Discriminant Function Coefficient

	Function
	1
How many members does your family have?	.939
According to the gratuated specialization	.580

Based on the standardized coefficients the correlation between the discriminant function and the independent variables is calculated, which is ranked by size of those coefficients [26]

It can be seen from these results (table no. 7) that the number of family members has a greater discrimination power than the other independent variable, because of the strongest association with the discriminant function.

Table no. 8 Correlations between independent variables and the discriminant function

Structure Matrix

	Function
	1
How many members does your family have?	.823
According to the gratuated specialization	.392

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

The system also provides the average scores of the groups based on the square of the distances as compared with the centroids group [27](table no.8).

Based on these scores the cutting score is calculated representing the average of the two scores in the table no.9 weighted by the size of the groups.

Thus, the separation score is determined by the formula:

$$CS = \frac{0.962 \cdot 39}{-0.879 \cdot 43} = 1$$

[26] Cristinel C. - Marketing information system. Marketing data analysis and processing. Applications in SPSS, Publisher Infomarket, Braşov, 2006, p. 244

[27] Malhotra N – Marketing research, Akadémia Press, Budapest, 2005, p. 653

Table no. 9 Average scores for both groups

Functions at Group Centroids

	Function
Do you have a job?	1
No	.969
Yes	-.879

Unstandardized canonical discriminant functions evaluated at group means

Depending on this score we can classify the respondents forming the two groups:

- if a respondent has the discriminatory score below zero, then it belongs to the group people who are unemployed since that group has negative discriminatory scores.
- if a respondent has the discriminatory score greater than zero, it will belong to the group of people who have a job.

The discriminatory score means the value of the discriminant function obtained by multiplying the coefficients of the discriminant function to the values of the independent variables of a person and adding them to the constant value

Based on the discriminatory scores the groups to which each respondent should be part of are determined. Finally, a centralization of the percentage of the persons belonging to the predicted group results, in terms of existence or lack of a job.

Table no. 10 Summary of the classification of respondents according to the predicted groups

Classification Results[b,c]

			Predicted Group Membership		
		Do you have a job?	No	Yes	Total
Original	Count	No	36	3	39
		Yes	7	36	43
	%	No	92.3	7.7	100.0
		Yes	16.3	83.7	100.0
Cross-validated[a]	Count	No	36	3	39
		Yes	7	36	43
	%	No	92.3	7.7	100.0
		Yes	16.3	83.7	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 87.8% of original grouped cases correctly classified.

c. 87.8% of cross-validated grouped cases correctly classified.

In the table no. 10 it is noted that in the case of the respondents who do not have a job, following the classification, 3 of respondents ought to be part of the group of those who have a jobs. Therefore, from the 39 respondents of the „original" analyzed group who were not employed 36 respondents (92.6%) are classified correctly with the

„predicted" group, while 3 of the respondents, (7.7%) ought to be a part of the group who has jobs if we were to take into account the discriminatory score.

In the case of the other group, 7 cases (16.3%) do not meet the classification based on the discriminatory score. These respondents, according to the classification, would not have to have a job, but they have it for real.

In the end, the conclusion is that a percentage of 87.8% of the cases correctly classified, the results of a cross validation of the model showing the same structure of the classification.

REFERENCES

1. Cristinel, C. Marketing information system. Marketing data analysis and processing. Applications in SPSS, Infomarket Press, Braşov, 2006
2. Malhotra, N Marketing Research, Akadémia Press, Budapest, 2005